

Dr Zenon Sosnowski
Wydział Informatyki
Politechnika Białostocka
ul. Wiejska 45a
15-351 Białystok

Białystok, 26.06.2013 r.

AUTOREFERAT

**przedstawiający opis dorobku i osiągnięć naukowych,
w szczególności określonych w art. 16 ust. 2 ustawy
(dotyczy okresu po uzyskaniu stopnia doktora)**

Posiadane dyplomy i stopnie naukowe

Doktor nauk technicznych

Politechnika Warszawska, Wydział Elektryczny; stopień uzyskany w czerwcu 1986 r.
Tytuł rozprawy doktorskiej: „Cyfrowa regulacja procesami nie w pełni zdeterminowanymi”.

Magister

Uniwersytet Warszawski, Wydział Matematyki i Mechaniki; (sekcja informatyki, specjalność języki i systemy programowania) dyplom z wynikiem bardzo dobrym uzyskany we wrześniu 1976 r. Tytuł pracy magisterskiej: „Badanie funkcji EVALQUOTE w LISP 1.5.9.”.

Informacje o dotychczasowym zatrudnieniu

Adiunkt

Politechnika Białostocka, Wydział Informatyki, Katedra Oprogramowania, IX/2010 – nadal

Z-ca Kierownika Katedry Oprogramowania

Politechnika Białostocka, Wydział Informatyki, IX/2008 – VIII/2010

Prodzian ds. Nauki Wydziału Informatyki

Politechnika Białostocka, IX/2002 – VIII/2008

Kierownik Zakładu Systemów Aplikacyjnych

Politechnika Białostocka, Wydział Informatyki, Katedra Oprogramowania, IX/2001 – VIII/2002

Adiunkt

Politechnika Białostocka, Instytut Informatyki, III/1998 – VIII/2001

Z-ca Dyrektora ds. Nauki Instytutu Informatyki (na prawach wydziału)

Politechnika Białostocka, I/1992 – II/1998

Visiting Fellow

Knowledge System Laboratory, Institute for Information Technology, National Research Council Canada, Ottawa, X/1989 – XII/1991

Kierownik Zakładu Informatyki

Politechnika Białostocka, III/1988 – IX/1989

Visiting Scientist

Delft University of Technology, the Netherlands, X/1987 – II/1988

Adiunkt

Politechnika Białostocka, Zakład Informatyki, X/1986 – IX/1987

Asystent

Politechnika Białostocka, Ośrodek ETO, X/1976 – IX/1986

Osiągnięcie naukowe

Jednotematyczny cykl publikacji pt.

„Rozmyte drzewa decyzyjne w analizie danych, generacji wiedzy i prognozowaniu”

- [1] W. Pedrycz (40%), Z.A. Sosnowski (60%), „Designing decision trees with the use of fuzzy granulation”, **IEEE Transactions on Systems, Man and Cybernetics, Part A** Vol. 30, No 2, 2000, str. 151-159.
JCR 35 pkt., IF / 5yearIF: 0,557 / 2,144, 30 cytowań.
- [2] W. Pedrycz (40%), Z.A. Sosnowski (60%), „The designing of decision trees in the framework of granular data and their application to software quality models”, **Fuzzy Sets & Systems**, vol. 124, (2001), p. 271-290.
JCR 40 pkt., IF / 5yearIF: 0,470 / 1,988, 19 cytowań.
- [3] W. Pedrycz (40%), Z.A. Sosnowski (60%), „C-Fuzzy Decision Trees”, **IEEE Transactions on Systems, Man and Cybernetics, Part C**, Vol. 35, No 4, 2005, p. 498-511. *JCR 30 pkt., IF / 5yearIF: 0,706 / 2,397, 19 cytowań.*
- [4] W. Pedrycz (40%), Z.A. Sosnowski (60%), „Genetically Optimized Fuzzy Decision Trees”, **IEEE Transactions on Systems, Man and Cybernetics, Part B**, Vol. 35, No 3, 2005, p. 633-41. *JCR 45 pkt., IF / 5yearIF: 1,108 / 3,337, 15 cytowań.*
- [5] Z.A. Sosnowski, „Activation of Fuzzy Rules in RETE Network”, w **„Flexible Query Answering Systems”**, H.L. Larsen et. al. (Eds.), Physica-Verlag/Springer-Verlag, 2000, str. 200-209.
- [6] Z.A. Sosnowski, „Comments on "FuzzyShell: A Large-Scale Expert System Shell Using Fuzzy Logic for Uncertainty Reasoning””, **IEEE Transactions on Fuzzy Systems**, Vol. 8, No 6, 2000, p.817-21.
JCR 45 pkt., IF / 5yearIF: 1,873 / 4,196, 2 cytowania.
- [7] Z.A. Sosnowski, „Chaining of fuzzy rules in expert systems”, **Lecture Notes in Computer Science** 2206, B. Reusch (Ed.), Springer-Verlag Berlin-Heidelberg, 2001, p. 895-903. *IF 0,415.*

Przy publikacjach, w nawiasach podany jest procentowy wkład w ich autorstwo. Dodatkowo podana jest liczba punktów przydzielona im zgodnie z aktualnym wykazem Ministerstwa Nauki i Szkolnictwa Wyższego, IF (zgodny z rokiem wydania), 5-letni IF oraz liczba cytowań (bez autocytaowań) g. bazy Web of Science.

Przedmiotem przeprowadzonych badań, których rezultaty zaprezentowano w postaci przedstawianego do oceny jednotematycznego cyklu publikacji, jest zastosowanie technik zbiorów rozmytych, jako narzędzia wspomagającego projektowanie drzew decyzyjnych.

Drzewa decyzyjne, będące powszechnie stosowaną architekturą uczenia maszynowego i systemów klasyfikacji, występują w literaturze z szeroką listą różnych schematów uczenia i upraszczania, różnorodnością algorytmów dyskretyzacji i szeregiem bardzo szczegółowych i wyrafinowanych ulepszeń. Pomimo takiej różnorodności będącej podstawą rozwoju drzew decyzyjnych, możemy łatwo zaobserwować kilka zasadniczych właściwości, które występują najwyraźniej w całym spektrum drzew decyzyjnych. Po pierwsze, drzewa działają na atrybutach dyskretnych, które przyjmują ograniczoną (zwykle raczej małą) liczbę wartości. Po drugie, na każdym etapie konstrukcji wybierany jest tylko jeden atrybut. Po trzecie, drzewa decyzyjne w ich ogólnej wersji są przeważnie stosowane do problemów dyskretnych klas (ciągłe problemy prognozy są obsługiwane przez drzewa regresyjne). Mając te wady wyraźnie zidentyfikowane, celem przeprowadzonych badań było zdefiniowanie nowej klasy drzew decyzyjnych, które spróbują złagodzić wspomniane wyżej problemy.

Publikacje [1] i [2] przedstawiają zastosowanie technik zbiorów rozmytych jako narzędzia wspomagającego projektowanie drzew decyzyjnych. W szczególności [1], skoncentrowano się na roli technik zbiorów rozmytych w problemie dyskretyzacji ciągłych atrybutów i pokazano, że rozmyte grupowanie kontekstowe może być efektywnie użyte do kwantyzacji atrybutów drzewa decyzyjnego. Tym nie mniej podkreślić należy, że w tej pracy rola teorii zbiorów rozmytych jest istotna tylko na etapie dyskretyzacji atrybutów, który poprzedza generowanie drzewa decyzyjnego. W tym sensie proponowana w [1] metoda zmniejsza dystans pomiędzy standardowymi technikami generowania drzew decyzyjnych (szczególnie dotyczącymi zadań kwantyzacji) a bardziej "radykalnymi" ujęciami prowadzącymi do koncepcji rozmytych drzew decyzyjnych, w których zastrzega się, że drzewo decyzyjne w swojej finalnej postaci wykonuje (w pewnym sensie) przetwarzanie informacji na poziomie zbiorów rozmytych.

Głównym zadaniem pracy [1] było pokazanie, że zbiory rozmyte mogą uważane jako użyteczne środowisko generowania drzew decyzyjnych dających w konsekwencji reguły decyzyjne oraz w jaki sposób środowisko to można efektywnie zrealizować. W publikacji podsumowano podstawowe koncepcje drzew decyzyjnych oraz objaśniono istotę proponowanych rozszerzeń. Praca miała na celu pomoc w identyfikacji niektórych problemów generowania drzew decyzyjnych. Przywołano ideę rozmytego grupowania kontekstowego i pokazano sposób, w którym to podejście staje się kluczowe w szerszym potraktowaniu ciągłych atrybutów. Praca pokazuje, że ta forma grupowania pomaga projektantowi na interakcje z potencjalnym użytkownikiem w zakresie przyjęcia określonych założeń dotyczących istotności zbioru danych. Szczegółowo omówiono dwa następujące problemy budowania drzewa decyzyjnego: sposoby definiowania kontekstowych skupień, dla których wykonywanie jest rozmyte grupowanie oraz metody dyskretyzacji (kwantyzacji) ciągłych atrybutów. Wyczerpująco przedstawiono wyniki eksperymentów obliczeniowych.

Drzewa decyzyjne są powszechnie spotykane w różnych zadaniach związanych z klasyfikacją. Ich główne zalety to przejrzystości i czytelności konstrukcji umożliwiające łatwe uzyskanie zbioru reguł decyzyjnych. Podczas przeglądu drzewa decyzyjnego spotykamy kolejne atrybuty zorganizowane według ich znaczenia w sensie zdolności do istotnego podziału zbioru danych. Cechy te powodują, że drzewa decyzyjne są dobrym narzędziem budowania złożonych modeli, takich jak np. model jakości oprogramowania wyrażonego poprzez różne jego mierniki.

Praca [2] przedstawia szczegółowo algorytm kontekstowego grupowania rozmytego oraz jego zastosowanie do budowy modelu jakości oprogramowania zrealizowanego w postaci drzew decyzyjnych. Pokazano, jak drzewa decyzyjne pomagają uchwycić podstawowe relacje pomiędzy zbiorem wskaźników oprogramowania w odniesieniu do wydajności poszczególnych modułów (wyrażonej np. liczbą modyfikacji kodu źródłowego).

W pracy [3], wprowadzono pojęcie i opisano budowę rozmytych drzew decyzyjnych opartych na granulach (ziarnach) informacji - wielowymiarowych encjach charakteryzujących się wysoką homogenicznością oraz niską zmiennością. Ponieważ takie granule są tworzone przez grupowanie rozmyte oraz grają kluczową rolę w generowaniu drzew decyzyjnych, nazywano takie drzewa klastrowymi rozmytymi drzewami decyzyjnymi (w skrócie *K-RDD*). Przyjmując pojęcia klastrowych (i rozmytego grupowania), możemy uwzględnić ciągłą naturę klas tak, że nie ma żadnego ograniczenia użycia *K-RDD* podobnie jak do problemów dyskretnych. W przeciwieństwie do "standardowych" drzew decyzyjnych, w których rozważamy tylko jedną zmienną (cechę) jednocześnie, ta forma drzew decyzyjnych wiąże za sobą wszystkie zmienne, które są rozważone w każdym węźle drzewa. Oczywiście, daje to początek do zupełnie nowej geometrii podziału przestrzeni cechy, która jest zupełnie inna od podziału generowanego przez standardowe drzewa decyzyjne. Generowanie klastrowego rozmytego drzewa decyzyjnego jest realizowane przez rozszerzanie węzła drzewa charakteryzującego się najwyższą zmiennością znajdującej się tam granuli informacji. Pokazano funkcjonalne i algorytmiczne szczegóły sposobu generowania drzewa w zależności od kilku dodatkowych kryteriów rozszerzenia węzła takich jak moc (liczba danych) w danym węźle i poziomie strukturalnych zależności istniejących w tym miejscu danych. Następnie, omówiono dokładnie działanie *K-RDD* w trybie klasyfikacji albo prognozowania gdy w ogólności dopuszczając ciągle wartości przyjmowane przez wyjście mamy do czynienia z ciągłym problemem prognozy. Klastrowe rozmyte drzewo decyzyjne może też być widziane jak pewien szkielet dla kilku szczegółowych modeli, które mogą być tworzone na bazie lokalnej poprzez rozważanie danych alokowanych w indywidualnych węzłach. Na tym etapie, możemy poddać obróbce modele przez wybieranie ich topologii (na przykład: modele liniowe, sieci neuronowe, itd.) oraz podjąć decyzję o szczegółach procesu uczenia. W końcowej części publikacji [3] przedstawiona jest seria eksperymentów obliczeniowych z wykorzystaniem zarówno danych syntetycznych jak i zestawów danych z repozytorium uczenia się maszynowego. Uzyskane wyniki są porównane z wynikami wyprodukowanymi przez "standardową" wersję drzewa decyzyjnego (mianowicie *C4.5*).

Praca [4], przedstawia koncepcję genetycznie optymalizowanych rozmytych drzew decyzyjnych (w skrócie *G-RDD*). Pojęciowa prostota i przejrzystość są to dwa oczywiste atuty, które odgrywają podstawową rolę w szerokiej popularności drzew decyzyjnych. Ta atrakcyjność ogólnej konstrukcji ostatecznie pociąga za sobą szereg różnych ulepszeń drzew decyzyjnych. Występuje to na poziomie nowych architektur, paradygmatów uczenia i mechanizmów interpretacji. Pojęciowe ramy Obliczeń Inteligentnych (w skrócie *OI*) powstają tutaj jako wszechstronny kierunek, w granicach którego drzewa decyzyjne mogą być ulepszone co do ich możliwości reprezentacji wiedzy. Rozwój różnych strategii uczenia jest też dobrze wspierany przez techniki *OI*. Uwzględniając duży potencjał obliczeń ewolucyjnych (algorytmów genetycznych, w szczególności), celem było zaprojektowanie rozmytych drzew decyzyjnych. Zastosowawszy kombinację dwu dominujących technik *OI*

takich jak zbiory rozmyte i metody ewolucyjne, możemy odnosić się do powstałej konstrukcji jako do *genetycznie* optymalizowanych *rozmytych* drzew decyzyjnych. W pracy [4] przedstawiono dwufazową konstrukcję *G-RDD*. Punktem wyjścia procesu generowania *G-RDD* jest drzewo z dyskretnymi albo przedziałowymi wartościami atrybutów, którego tworzenie oparte jest na "standardowej" technice takiej jak *C4.5*. Jest kilka powodów stojących za tym dość tradycyjnym sposobem projektowania drzewa. Algorytm *C4.5* jest dobrze udokumentowany i ulepszony na kilka sposobów. Jego praktyczna użyteczność jest widoczna ponieważ doczekała się wielkiej liczby opracowań naukowych. Z punktu widzenia optymalizacji, *C4.5* jest całkiem prosty i niezbyt wymagający od strony obliczeniowej. W skrócie, *C4.5* jest iteracyjną techniką rozszerzania drzewa w tym sensie, że wybieramy "najlepszy" atrybut lokalnie (mianowicie zlokalizowany przy określonym węźle drzewa) i taki proces wyboru występuje w każdym kroku procedury optymalizacji. Ta strategia nie uwzględnia żadnego nawracania aby wspomagać ewentualne podnoszenie jakości i poprawę zrealizowanej dotychczas struktury. Optymalizowanie drzewa w każdym węźle nie koniecznie znaczy, że uzyskamy drzewo, które jest optymalne w globalnym sensie. Druga faza, w której drzewo jest ulepszone, przez włączanie zbiorów rozmytych połączonych z indywidualnymi węzłami, ma na celu w optymalizację (podnoszenie jakości) budowanej struktury. Cel ten jest realizowany przez użycie optymalizacji ewolucyjnej, w szczególności algorytmów genetycznych. W tym sensie, możemy powiedzieć o współdziałaniu techniki zbiorów rozmytych i obliczeń ewolucyjnych. Zbiory rozmyte podnoszą zdolności drzewa do reprezentacji wiedzy. Algorytmy genetyczne pomagają w wykonaniu optymalizacji drzewa. Ze swej natury, możemy skorzystać na ich zdolności do wykonania w pewnym sensie globalnego wyszukiwania. Bardziej szczegółowo, uogólniamy drzewo decyzyjne do jego rozmytego odpowiednika przez dopuszczenie zbiorów rozmytych jak komponentów decyzji ustawionych w indywidualnych węzłach. Następnie, przedstawiono wszechstronną i systematyczną dwufazową procedurę konstrukcji, która zaczyna się od klasycznego *C4.5* i następnie poprawia wynik w ramach zastosowania algorytmów genetycznych. To co powstaje, jest odzwierciedlone w samej nazwie konstrukcji, która brzmi genetycznie optymalizowane rozmyte drzewo decyzyjne. Artykuł [4] omawia dokładnie kształtowanie i użycie zbiorów rozmytych jako pojedynczych wartości atrybutów oraz skupia się na funkcjonalnym modelu zaproponowanego drzewa decyzyjnego, w szczególności omawia jak wspomagać mechanizmy propagacji wykonywane w granicach drzewa oraz jak przenieść agregację wyników zebranych w węzłach końcowych aby znaleźć ogólną odpowiedź drzewa dla danego wejścia. Rozważone są zarówno zagadnienia w formacie dyskretnym jak i ciągłym. Szereg eksperymentów obliczeniowych wykorzystujących dane syntetyczne i dane z repozytorium uczenia maszynowego zostało użytych by zilustrować wydajność proponowanych procedur *G-RDD*.

Proponowane w pracach [1-4] algorytmy generują rozmyte drzewa decyzyjne, które łatwo zinterpretować jako zbiór reguł decyzyjnych a następnie wykorzystać, w standardowy sposób, do utworzenia rozmytego systemu ekspertowego. Architektura takiego rozmytego systemu ekspertowego została przedstawiona w pracach [5-7].

Wiele informacji znajdujących się w bazie wiedzy systemu ekspertowego jest nieprecyzyjna, niekompletna lub nie całkowicie wiarygodna. W publikacji [5] opisano zaproponowaną architekturę rozmytego systemu ekspertowego. Prezentowany system powstał na bazie rozszerzenia architektury szkieletowego systemu ekspertowego opartego o sieć *RETE*. Sieć

RETE w efektywny sposób realizuje centralny algorytm systemu regułowego, którym jest dopasowywanie bieżących danych do przesłanek reguł. Rozszerzony rozmyty system ekspertowy zawiera możliwości wykorzystania koncepcji zbiorów rozmytych w reprezentacji wiedzy i mechanizmach wnioskowania. Rozszerzona w ten sposób funkcjonalność tradycyjnego systemu ekspertowego umożliwia wykorzystanie wnioskowania boolowskiego i rozmytego, oraz ich kombinacje, pozwalając na dowolne łączenie terminów rozmytych i nierozmytych w bazie wiedzy systemu ekspertowego. Umożliwia to ekspertowi dziedzinowemu wyrażenie reguł w swoich własnych terminach rozmytych. System wykorzystuje dwie podstawowe koncepcje niepewności: rozmytość i wiarygodność. Reprezentacja rozmytości oparta jest na koncepcji zmiennej lingwistycznej, tj. zmiennej, która przyjmuje wartości nie jako liczby ale jako słowa albo zdania języka naturalnego albo formalnego. Z każdą zmienną lingwistyczną skojarzona jest odpowiednia składnia i semantyka. Składnia generuje zbiór słów lub wyrażeń opisujących wszystkie możliwe wartości zmiennej lingwistycznej. Semantyka dostarcza unikalne wartości rozmyte ze zbioru potęgowego przestrzeni rozważań zmiennej lingwistycznej dla każdego składniowo poprawnego wyrażenia. Wiarygodność pojawia się wtedy, kiedy nie ma absolutnej pewności co do prawdziwości pojedynczej informacji. Stopień wiarygodności jest reprezentowany przez liczbę rozmytą lub nierozmytą. W procesie wnioskowania wykorzystuje się logikę rozmytą i liczby rozmyte. Zbiory i relacje rozmyte modelują aspekty wnioskowania rozmytego we wnioskowaniu przybliżonym podczas gdy liczby rozmyte wykorzystywane są do przetwarzania współczynników wiarygodności.

Zaproponowana w [5] analiza reguł jest realizowana przez rozszerzoną sieć *RETE* i zależy od tego, czy poprzedniki reguł (*LHS*) lub ich następniki (*RHS*) zawierają obiekty rozmyte. Możemy wyróżnić trzy typy reguł: *CRISP__*, *FUZZY__CRISP* oraz *FUZZY__FUZZY*. Jeżeli poprzednik reguły nie zawiera obiektów rozmytych, wtedy reguła jest typu *CRISP_* bez względu na to czy następnik reguły zawiera obiekt rozmyty czy też nie. Gdy tylko poprzednik zawiera obiekt rozmyty, wtedy typ reguły jest *FUZZY_CRISP*. Jeżeli zarówno poprzednik jak i następnik zawierają obiekty rozmyte, wtedy reguła jest typu *FUZZY_FUZZY*. Jeżeli reguła jest typu *CRISP__*, wówczas odpowiada to tradycyjnemu systemowi ekspertowemu i stosowana jest klasyczna reguła wnioskowania *modus ponens*. Jeżeli reguła jest typu *FUZZY_CRISP*, wówczas współczynnik wiarygodności konkluzji jest modyfikowany przez miarę podobieństwa pomiędzy faktem aktywującym regułę a jej przesłanką. Miara podobieństwa obliczana jest na bazie miary możliwości i miary konieczności. W przypadku gdy reguła jest typu *FUZZY_FUZZY*, stosowana jest *uogólniona reguła wnioskowania modus ponens* a obliczenie konkluzji oparte jest na złożeniowej regule wnioskowania. Dla zbioru rozmytego, który jest wynikiem wnioskowania rozmytego, odpowiednia wartość zmiennej lingwistycznej może być obliczona na podstawie aproksymacji lingwistycznej. Jako wynik aproksymacji lingwistycznej, użytkownicy odnoszą się do wyrażeń lingwistycznych, podczas gdy proces wnioskowania jest wykonany z wykorzystywaniem obliczeń na zbiorach rozmytych. W przypadku kiedy ze zbioru rozmytego musi być wybrany pojedynczy element, ma miejsce proces wyostrzania.

W nierozmytych systemach ekspertowych fakty albo istnieją albo nie istnieją (w niektórych systemach za istniejące uważa się te fakty, których współczynniki wiarygodności są większe niż wartość progowa). W rozmytych systemach ekspertowych fakty zawierające obiekty rozmyte nie są jednoznacznie określane jak istniejące albo nie. W konsekwencji, fakt

zawierający obiekt rozmyty jest potencjalnie użyteczny dla każdej reguły, której poprzednik (LHS) zawiera ten obiekt rozmyty i każda wartość obiektu rozmytego może być wielokrotnie ponownie wyznaczana. Zarówno w systemach ekspertowych nierozmytych jak i rozmytych fakty mogą być dołączane na dwa sposoby. Albo użytkownik systemu (lub urządzenie zewnętrzne) może fakt dołączyć, albo baza wiedzy może go wydedukować. W tradycyjnym systemie ekspertowym nie pojawia się nigdy jakkolwiek potrzeba ponownego ustalania istnienia faktów w systemie - raz dołączone istnieją. Ale, jak zasugerowano powyżej, w systemie rozmytym, uszczegółowianie może być możliwe. Tak więc, w przypadku w którym fakt zawierający obiekt rozmyty jest dołączany jako wynik wykonywania prawej strony (RHS) reguły, ten fakt jest traktowany jako dający wsparcie potwierdzające w odniesieniu do konkluzji zawierających ten sam obiekt rozmyty. W pracy [6] a następnie bardziej szczegółowo w pracy [7] został zaprezentowany algorytm dynamicznego dodawania nowych faktów reprezentujących obiekty rozmyte do rozmytej sieci RETE w sposób, który wykorzystuje wykonane wcześniej testy i wyniki dopasowań. Zaproponowany algorytm efektywnie realizuje proces wielokrotnego dopasowywania faktów do przesłanek reguł, który jest najbardziej czasochłonnym cyklem w działaniu maszyny wnioskującej systemu regułowego.

Podsumowując, w ramach przeprowadzonych badań przedstawionych w jednotematycznym cyklu publikacji pt. „Rozmyte drzewa decyzyjne w analizie danych, generacji wiedzy i prognozowaniu” pokazano, że zbiory rozmyte mogą być uważane, jako użyteczne środowisko generowania drzew decyzyjnych dających w konsekwencji reguły decyzyjne. Pokazano także w jaki sposób można efektywnie zrealizować generowanie proponowanych rozmytych drzew decyzyjnych. W przedstawionych publikacjach, skoncentrowano się w szczególności na roli technik zbiorów rozmytych w problemie dyskretyzacji ciągłych atrybutów i pokazano, że rozmyte grupowanie kontekstowe [1] może być efektywnie użyte do kwantyzacji atrybutów drzewa decyzyjnego. Opracowane algorytmy pozwalają na optymalizację dyskretyzacji, która będąc kierowaną przez błąd klasyfikacji może być zrealizowana zarówno w trakcie jak i po sfinalizowaniu procesu generowania drzewa. Ponadto opracowane techniki granulacji [2,3,4], w tym m.in. granulacji kontekstowej, pozwalają uwzględnić ciągłą naturę klas tak, że nie ma żadnego ograniczenia ich użycia podobnie jak do problemów dyskretnych. Proponowane algorytmy nie wymagają skomplikowanych obliczeń co umożliwia stosowanie ich również w tworzeniu reguł decyzyjnych. Powstały na bazie zoptymalizowanego rozmytego drzewa decyzyjnego zbiór reguł decyzyjnych, może być następnie wykorzystany, w standardowy sposób, do utworzenia rozmytego systemu ekspertowego. W pracach [5,6,7] przedstawiłem i wszechstronnie przebadalem architekturę rozmytego systemu ekspertowego opartego na zaproponowanym w [5] rozszerzeniu sieci RETE.

Omówienie pozostałych osiągnięć naukowo-badawczych

Główne wyniki moich prac naukowych w pierwszych latach po uzyskaniu stopnia doktora zostały zebrane w dwu artykułach opublikowanych w prestiżowych czasopismach *Fuzzy Sets and Systems* [8] oraz *Computers & Artificial Intelligence* [9]. W pracach tych został zaproponowany nowy język programowania (zwany FLISP), który zapewnia szereg udogodnień dla skutecznego reprezentowania i przetwarzania danych rozmytych.

Zdefiniowany jako rozszerzenie standardowego języka LISP, oparte na logice rozmytej i teorii zbiorów rozmytych. Język zawiera zbiór operacji na zbiorach rozmytych oraz procedury dotyczące rozwiązań rozmytych równań relacyjnych z normami trójkątnych. FLISP stanowi wygodne narzędzie do realizacji i badania algorytmów sterowania rozmytego. W pracy [8] opisano składnię, struktury danych oraz instrukcje języka FLISP.

Przykład wykorzystania języka FLISP jako funkcjonalnego narzędzia specyfikacji oprogramowania wykonanego z pomocą tabel decyzyjnych jest przedstawiony w pracy [10]. Przedstawiona propozycja oznacza kolejne spojrzenie na regulator rozmyty jako oprogramowanie, które jest całkowicie opisane przez jego tabele decyzyjne z wpisami utworzonymi w postaci zbiorów rozmytych.

W pracach [11, 12] zaproponowano rozszerzenie paradygmatu obiektowego modelu danych z wykorzystaniem zbiorów rozmytych. W szczególności zostało zdefiniowanie kluczowe pojęcie obiektowego modelu danych tj. rozmyte dziedziczenie z wykorzystaniem aproksymacji lingwistycznej oraz metoda klasyfikacji obiektów w hierarchii klas rozmytych. Rozmyta sieć neuronowa została wykorzystana na poziomie uczenia i agregacji wyników dopasowania kolejnych atrybutów. Przydatność proponowanego podejścia zilustrowano przykładami analizy scen [11], zwłaszcza klasyfikacji obiektów statycznych i ruchomych oraz problemem klasyfikacji pojazdów [12].

W pracach [13, 14] zaproponowano użycie teorii Dempstera-Shafera jako narzędzia wspomagającego projektowanie rozmytych reguł decyzyjnych. Prace te przedstawiają zastosowanie rozmytego klasyfikatora Dempstera-Shafera. W szczególności, koncentrowano się na roli operatorów rozmytych i na problemie dyskretyzacji ciągłych atrybutów. Pokazano, w jaki sposób proponowane metody mogą być skutecznie wykorzystywane w kwantyzacji atrybutów dla generacji reguł rozmytych. Praca [13] analizuje problem wpływu doboru operatorów rozmytych na jakość otrzymywanego zbioru reguł. W pracy [14] badano problem dostrajania parametrów funkcji przynależności zmiennych lingwistycznych z wykorzystaniem techniki algorytmów genetycznych. Wyczerpująco przedstawiono wyniki eksperymentów obliczeniowych wykonanych dla kilku zbiorów danych medycznych dostępnych w sieci Internet.

Praca [15] wprowadza koncepcję klasyfikatora opartego na binarnych diagramach decyzyjnych (BDD). Istota proponowanego podejścia leży w zastosowaniu algorytmu genetycznego do optymalnej binaryzacji ciągłych atrybutów dającej w konsekwencji binarny diagram decyzyjny. Jak wiadomo, forma i wielkość BDD jest bardzo wrażliwa na kolejności zmiennych decyzyjnych. W celu optymalizacji BDD zastosowano występujące w literaturze heurystyczne algorytmy wyznaczające kolejność zmiennych pozwalającą na minimalizację BDD. Powstały BDD pomagają kompresować dane, odkryć najbardziej istotne atrybuty zbioru danych oraz wykonać zadanie klasyfikacji.

W pracy [16] przedstawiono zastosowanie podsumowania lingwistycznego jako predykatu rozmytego do wyznaczania obiektów z typową wartością atrybutu lub zbioru atrybutów. W rozmytym algorytmie grupującym wykorzystana jest populacja z wyznaczoną ze względu na dany atrybut typowością obiektów. Wyniki działania tego algorytmu oraz jego zmodyfikowanej postaci zostały przedstawione na przykładzie populacji, której obiektami są piksele obrazu.

Praca [17] przedstawia modyfikację zaproponowanego w [16] algorytmu grupowania obiektów opartego na podsumowaniach lingwistycznych. Proponowana modyfikacja poprawia grupowanie ale powoduje wzrost i tak dużej w tym algorytmie złożoności. W celu przetwarzania dużej liczby danych zaproponowano algorytm równoległy. Algorytm ten został zaimplementowany korzystając ze standardu MPI do komunikacji między procesami działającymi na różnych procesorach. W pracy [17] przedstawiono algorytm równoległy i wyniki eksperymentów.

- [8] Z.A. Sosnowski, "FLISP - A Language for Processing Fuzzy Data", **Fuzzy Sets and Systems**, Vol.37, 1990, p.23-32.
JCR 40 pkt., IF / 5yearIF: 0,364 / 1,988, 9 cytowań.
- [9] Z.A. Sosnowski, "Data Structures for Representing and Processing of Fuzzy Information in LISP", **Computers & Artificial Intelligence (teraz Computing and Informatics)**, Vol. 10, 1991, No. 6, p. 561-71; NRCC Publication No. 33158.
JCR 15 pkt., IF : 0,181, 1 cytowanie.
- [10] W. Pedrycz (40%), Z.A. Sosnowski (60%), "FLISP: Representing and Processing of Uncertain Information", in **Fuzzy Logic for the Management of Uncertainty**, L.A. Zadeh and J. Kacprzyk (Eds.), John Wiley & Sons, Inc., New York, 1992, pp. 495-521.
- [11] W. Pedrycz (40%), Z.A. Sosnowski (60%), „Fuzzy Object-Oriented System Design”, **Fuzzy Sets & Systems**, Vol. 99, 1998, str. 121-134.
JCR 40 pkt., IF / 5yearIF: 0, 331 / 1,988, 4 cytowania.
- [12] W. Pedrycz (40%), Z.A. Sosnowski (60%), „FOOD: Towards Fuzzy Object-Oriented System Design”, w „**Computational Intelligence and Software Engineering**”, W. Pedrycz, J.F. Peters, (Eds.), World Scientific, 1999, p. 269-94.
- [13] Z.A. Sosnowski (50%), J.S. Walijewski (50%), „Generating Fuzzy Decision Rules with the Use of Dempster-Shafer Theory”, **13th European Simulation Multiconference**, June 1-4, 1999, Warsaw, Vol. II, p.419-25.
- [14] J.S. Walijewski (50%), Z.A. Sosnowski (50%), „Genetic Tuning Fuzzy Dempster-Shafer Decision Rules”, **TASK QUARTERLY**, Vol 6, No4 (2002), 631-640.
(6 pkt. Lista czasopism MNiSW).
- [15] W. Pedrycz (40%), Z.A. Sosnowski (60%) „Genetic BDD-based Pattern Classifier”, in **Information Processing and Security Systems**, K. Saeed and J. Pejaś(Eds), Springer, 2005, p. 257-268.
- [16] M.K. Smolińska (50%), Z.A. Sosnowski (50%), „Podsumowania lingwistyczne z grupowaniem rozmytym”, **Zeszyty Naukowe Politechniki Białostockiej. Informatyka**, 2007, Zeszyt 2, str. 141-154. *(6 pkt. Lista czasopism MNiSW).*
- [17] M.K. Smolińska (50%), Z.A. Sosnowski (50%), „Linguistic Summary with Parallel Fuzzy Clustering”, **Zeszyty Naukowe Politechniki Białostockiej. Informatyka**, 2009, Zeszyt 4, str. 139-150. *(6 pkt. Lista czasopism MNiSW).*

Jestem autorem lub współautorem 84 prac naukowych. Na ten mój całkowity dorobek naukowy składają się:

- 9 artykułów w czasopismach znajdujących się w bazie Journal Citation Reports,
- 5 artykułów w czasopismach spoza bazy Journal Citation Reports,
- 5 rozdziałów w monografiach w języku angielskim,
- 45 artykułów w materiałach konferencyjnych,
- 20 prac nieopublikowanych (sprawozdania z badań).

Sumaryczny Impact Factor moich publikacji naukowych według listy **Journal Citation Reports**, zgodnie z rokiem opublikowania wnosi **6,005**.

Liczba cytowań bez autocytowań któregośkolwiek z autorów według bazy **Web of Science** wynosi **97**.

Indeks Hirscha moich publikacji według bazy **Web of Science** wynosi **5**.

Brałem udział w następujących projektach badawczych:

- Temporalna reprezentacja wiedzy i jej implementacja w informatycznych systemach wspomaganie postępowania medycznego, grant MNiSW, wykonawca;
- Eksploracja wiedzy w bazach danych z wykorzystaniem koncepcji bionicznych, COST Action 282, główny wykonawca;
- 5 prac statutowych Wydziału Informatyki PB (kierownik – 2, główny wykonawca - 3);
- 6 prac własnych Politechniki Białostockiej (kierownik).

Byłem 4-krotnie nagradzany przez Rektora Politechniki Białostockiej za działalność naukową. Uzyskałem 5 międzynarodowych grantów konferencyjnych.

Wielokrotnie prezentowałem swoje wyniki na międzynarodowych i krajowych konferencjach naukowych oraz seminariach. Uczestniczyłem w 18 międzynarodowych konferencjach naukowych oraz w ponad 30 konferencjach krajowych.

Główne moje referaty znajdują się, przede wszystkim, w materiałach następujących konferencji:

- The European Simulation & Modelling Conference ESM'2007
- International Conference on „Fuzzy Sets Theory and its Applications”
- (IFSA) Seventh International Fuzzy Systems Association World Congress
- NATO Advanced Study Institute on „Soft Computing and Its Applications”
- 13th European Simulation Multiconference
- Joint EUROFUSE-SIC'99 International Conference
- Joint IFSA-EC and EURO-WG Workshop on Fuzzy Sets
- The International Fuzzy Engineering Symposium
- The International Conference "Fuzzy Sets in Informatics"
- Joint IFSA-EC and EURO-WG Workshop "Progress in Fuzzy Sets in Europe"
- NAFIPS – North America Fuzzy Information Processing Society Conference
- European Meeting on Cybernetics and Systems Research,
- International AMSE Conference "SYSTEMS: Analysis, Control & Design"
- Warsztaty Naukowe PTSK „Symulacja w Badaniach i Rozwoju”
- Konferencja „Inżynieria Wiedzy i Systemy Ekspertowe”
- Krajowa Konferencja Naukowa nt. „Sztuczna inteligencja”
- Ogólnopolskie Konwersatorium nt. „Sztuczna inteligencja”
- Krajowa Konferencja “Komputerowe Wspomaganie Badań Naukowych”

Omówienie dorobku dydaktycznego i popularyzatorskiego oraz informacje o współpracy międzynarodowej

Byłem przewodniczącym komitetu organizacyjnego kilku edycji konferencji „Technologie Eksploracji i Reprezentacji Wiedzy” oraz członkiem komitetu organizacyjnego 13th European Simulation Multiconference i kilku edycji Warsztatów Naukowych PTSK „Symulacja w Badaniach i Rozwoju”.

Jestem sekretarzem redakcji kwartalnika „Symulacja w Badaniach i Rozwoju”, które znajduje się w wykazie czasopism punktowanych na stronie Ministerstwa Nauki i Szkolnictwa Wyższego z liczbą punktów 2 za umieszczoną w nim publikacją. Wykonałem recenzje kilku artykułów spośród zaproponowanych do opublikowania w tym czasopiśmie.

Jestem członkiem komitetu redakcyjnego biuletynu informacyjnego EUROSIMu „Simulation News Europe”.

Jestem członkiem Polskiego Towarzystwa Informatycznego od 1982 roku. Od 2011 roku jestem członkiem Zarządu Głównego PTI, od 2011 r. prezesem Oddziału Podlaskiego PTI. W latach 2007-2011 byłem przewodniczącym Kola Podlaskiego PTI. Od 1993 roku jestem członkiem Zarządu Głównego Polskiego Towarzystwa Symulacji Komputerowej. Przez 3 kolejne kadencje byłem sekretarzem PTSK. Aktualnie po raz trzeci pełnię funkcję skarbnika PTSK.

W Politechnice Białostockiej, gdzie jestem zatrudniony, prowadziłem wykłady i pracownie specjalistyczne m.in. z podstaw programowania, systemów ekspertowych, systemów rozmytych oraz pracownie i seminaria dyplomowe. Byłem promotorem ponad 70 prac magisterskich, ponad 60 prac inżynierskich i ponad 60 prac licencjackich. Otrzymałem 13 nagród Rektora Politechniki Białostockiej w tym 4 indywidualne.

Odbyłem sześć krótkich staży naukowych w zagranicznych ośrodkach akademickich:

- Department of Electrical & Computer Engineering, University of Alberta, Edmonton, Canada, 2000, 2002, 2004;
- Department of Electrical & Computer Engineering, The University of Manitoba, Winnipeg, Canada, 1995, 1998;
- Department of Mathematics and Computer Science, Delft University of Technology, The Netherlands, 1983.

Recenzowałem ponad 50 zagranicznych projektów badawczych w tym m.in.: 11 projektów dla International Bureau of the Federal Ministry of Education and Research (Niemcy), 3 projekty dla programu SEE-ERA.NET PLUS oraz ponad 30 dla Komisji Europejskiej w ramach programów: eContentplus, INTAS, FP5.

Byłem recenzentem ponad 10 projektów krajowych w tym m.in.: 2 badawczych, 3 wdrożeniowych (POIG) oraz ponad 9 inwestycyjnych (ZPORR i POIG).

Recenzowałem ponad 50 publikacji dla czasopism znajdujących się w bazie Journal Citation Reports (JCR), m.in: IEEE Transactions on Fuzzy Systems, Fuzzy Sets and Systems, IEEE Transactions on Systems, Man, and Cybernetics, IEEE Transactions on Data and Knowledge Engineering, Mathematical and Computer Modelling of Dynamical Systems, International Journal of Approximate Reasoning, Information Sciences, Fuzzy Optimization and Decision Making, Pattern Analysis & Applications Journal, IEEE Transactions on Neural Networks, Journal of Multiple-Valued Logic and Soft Computing, Journal of Systems Science and System Engineering,

Byłem członkiem komitetów programowych kilku konferencji, m.in.: 2013 IFSA World Congress and NAFIPS Annual Meeting, 5th World Multi-Conference on Systemics, Cybernetics and Informatics, 13th European Simulation Multiconference, Information Technology Based Higher Education and Training (dwukrotnie), Warsztaty Naukowe PTSK „Symulacja w badaniach i rozwoju” (19 razy) oraz Konferencja „Technologie Eksploracji i Reprezentacji Wiedzy” (ośmiokrotnie).

Byłem odznaczony Srebrnym Krzyżem Zasługi (1999), Medalem Złotym za Długoletnią Służbę (2008) oraz Medalem Komisji Edukacji Narodowej (2005).

Zenon Grawoński